JOC The Journal of Organic Chemistry

Note

Modular Synthesis of Constrained Ethyl (cEt) Purine and Pyrimidine Nucleosides

Helen Blade, Derek Bradley, Louis Diorazio, Timothy Evans, Barry R. Hayter, and Gareth P Howell

J. Org. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.joc.5b00607 • Publication Date (Web): 20 Apr 2015

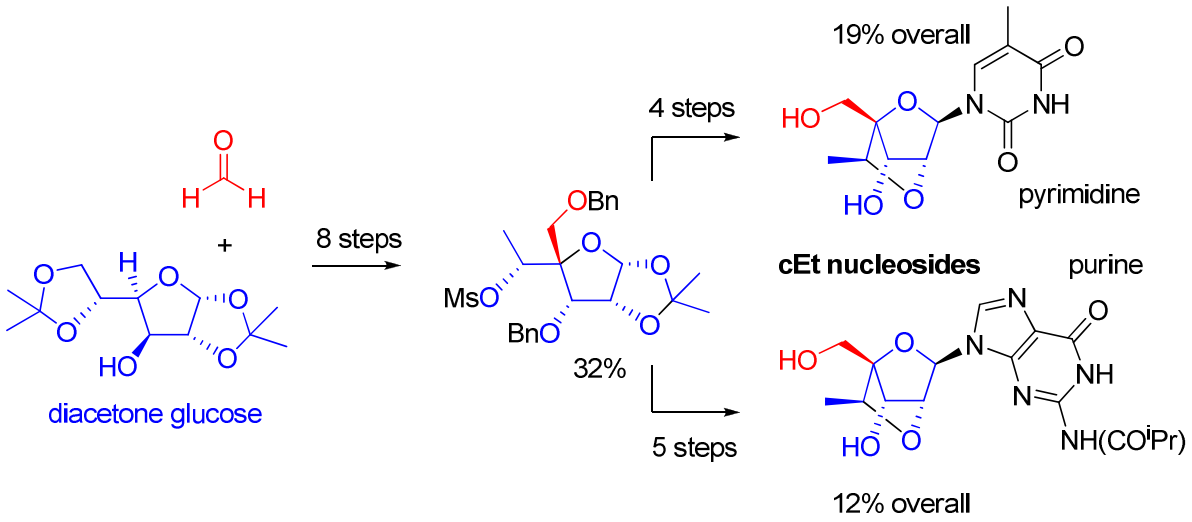
Modular Synthesis of Constrained Ethyl (cEt) Purine and Pyrimidine Nucleosides

Helen Blade, Derek Bradley, Louis Diorazio, Timothy Evans, Barry R. Hayter, Gareth P. Howell*

*Pharmaceutical Development, AstraZeneca, Charter Way, Macclesfield, SK10 2NA, United Kingdom

gareth.howell@astrazeneca.com

TOC Graphic

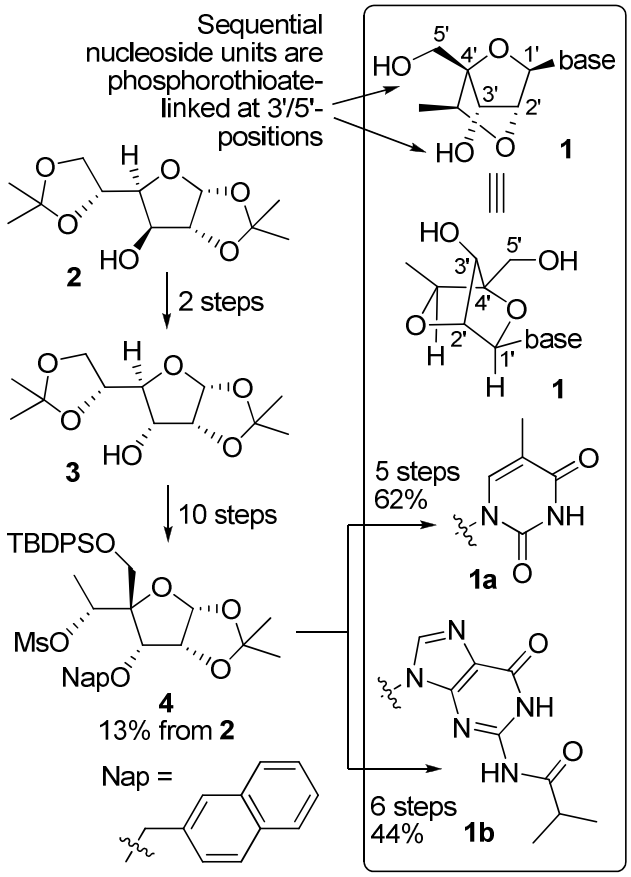


Abstract

A modular and scalable approach to pyrimidine- and purine-containing constrained ethyl (cEt) nucleosides is demonstrated. Minimising stereochemical adjustments and protecting group manipulations, diacetone glucose is converted to two representative cEt nucleosides via a functionalised, common intermediate. The retrosynthetic approach to this complex class of drug precursors offers clear benefits over existing routes based on step count and throughput/efficiency.

As newer technologies progress away from traditional DNA-based platforms, there is a growing interest in oligonucleotide therapeutics comprising novel nucleoside units.[1] One such motif, introduced by ISIS Pharmaceuticals,[2] utilises a [2.2.1] tricyclic core as exemplified by the pyrimidine (**1a**) and purine (**1b**) cEt nucleosides shown in **SCHEME 1**. A number of cEt-based oligonucleotides are currently undergoing clinical investigation.[3,4] The 4'-substitution in the furanose core of **1** poses a significant challenge,[5] and at 40+ steps, the synthetic burden of producing the (typically) four constituent cEt nucleosides for an oligonucleotide drug represents one of the largest seen in the "small molecule" arena.

SCHEME 1. Synthesis of cEt nucleosides

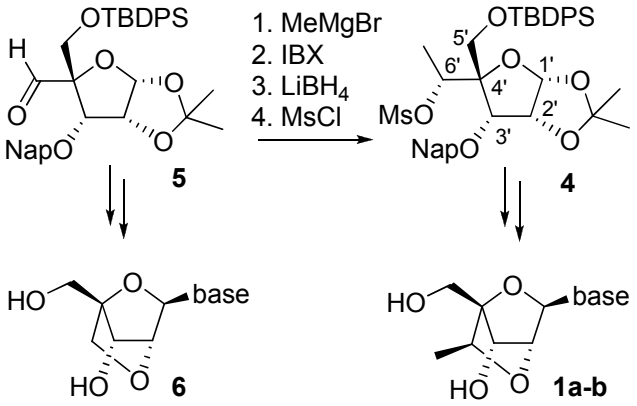


The synthesis of an individual cEt nucleoside such as **1a** can be achieved using short (10 steps from 5-methyluridine), linear synthetic sequences as reported recently by Hanessian et al.[6] For the large-scale manufacture of an oligonucleotide drug, however, multiple cEt nucleosides are required (as well as **1a** and **1b**, there are various other reported cEt nucleosides comprising uracil, cytosine, methyl cytosine and

adenine bases).[2,7,8,9,10] The summation of multiple 10-step linear sequences quickly becomes unwieldy for commercial use and a modular approach would be expected to offer reduced overall step-count and supply chain flexibility.

The only demonstrated route to multiple cEt nucleosides (**2→3→4→1**; see **SCHEME 1**) was reported in 2010 and has since been modified allowing access to multi-kilogram quantities of cEt nucleosides **1**.[2] In this current approach, diacetone allofuranose **3** (derived from diacetone glucose **2**) is converted to building block **4** in 13% yield over 12 steps and is subsequently functionalised to give cEt nucleosides **1**. Beginning in 2013, we have been investigating alternative synthetic routes to cEt nucleosides such as **1a** and **1b** with a focus on fewer overall steps (23 currently to produce **1a-b** from **2**),[11] and greater atom economy (60% of the molecular mass of mesylate **4** is contained in the three protecting groups). A further requirement for any new synthetic approach is the avoidance of constrained methyl (cMe) side-products of the type **6** shown in **SCHEME 2**. These cMe side-products, being generated from residual aldehyde **5**, are closely related to the desired cEt structures **1** and are persistent and difficult to remove. Avoidance of these side-products would represent a significant improvement with respect to eventual drug purity.[12]

SCHEME 2. Generation of cMe impurities

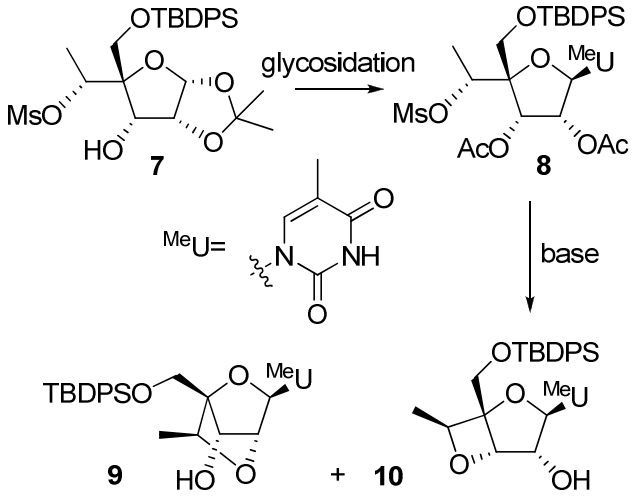


We decided to investigate analogs of the mesylate **4** with alternative protecting groups at the 3'- and 5'-positions. In keeping with results reported by Hanessian et al,[6] we found that glycosidation of structures such as **7** can be achieved in the absence of 3'-protecting groups as shown in **SCHEME 3**. Although this is attractive from an efficiency perspective, the subsequent cEt-forming cyclisation (**8→9**) becomes

markedly more complex with a competing oxetane-forming pathway leading to structures such as **10**. Due to our concern that the eventual cEt oligonucleotide drug could be contaminated with closely-related oxetane glycosides (representing a safety and efficacy risk)[12] we elected to continue with a dual, 3',5'-protecting group strategy using the same protecting group for both, such that they might be introduced and removed jointly, improving efficiency and minimising step count.

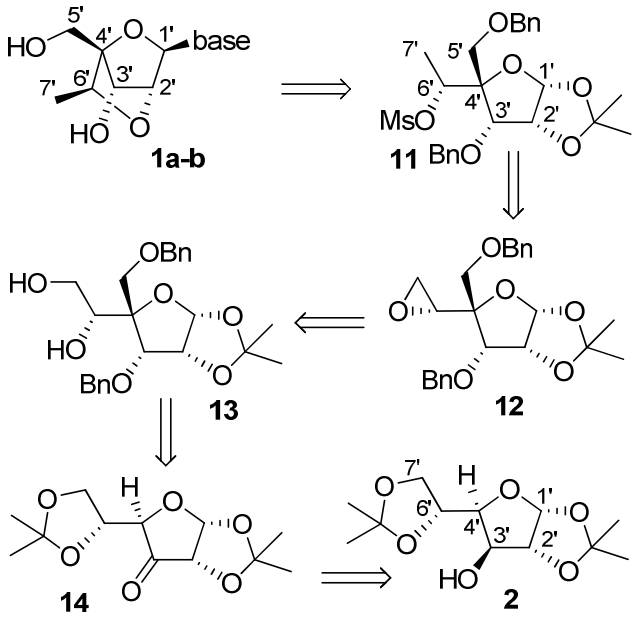
There are few protecting groups that are both (i) capable of withstanding the required glycosidation/cyclisation procedures [13] and (ii) amenable to large-scale use.[14] Benzyl (Bn) groups were selected with the expectation that hydrogenolysis could be carried out to complete the nucleoside synthesis,[15] avoiding the DDQ and HF deprotection procedures employed currently.[2]

SCHEME 3. Synthesis using one protecting group



This required a synthetic route to mesylate **11** and our retrosynthetic approach is shown in **SCHEME 4**. We were unable to identify a cheaper, more readily available starting material than diacetone glucose **2** and we envisaged that the pendant 6'(*R*)-mesylate in **11** could be derived from epoxide **12** via terminal hydride reduction and mesylation of the resulting secondary alcohol. Using this approach, the natural 6'-stereochemistry of glucose is utilised without the need for stereochemical adjustments.[16] Moreover, retention of the 7'-C atom from diacetone glucose **2** circumvents the homologation sequence used currently (**SCHEME 2**)[2] thus avoiding cMe impurities of the type **6** highlighted earlier.[12]

SCHEME 4. Retrosynthetic analysis

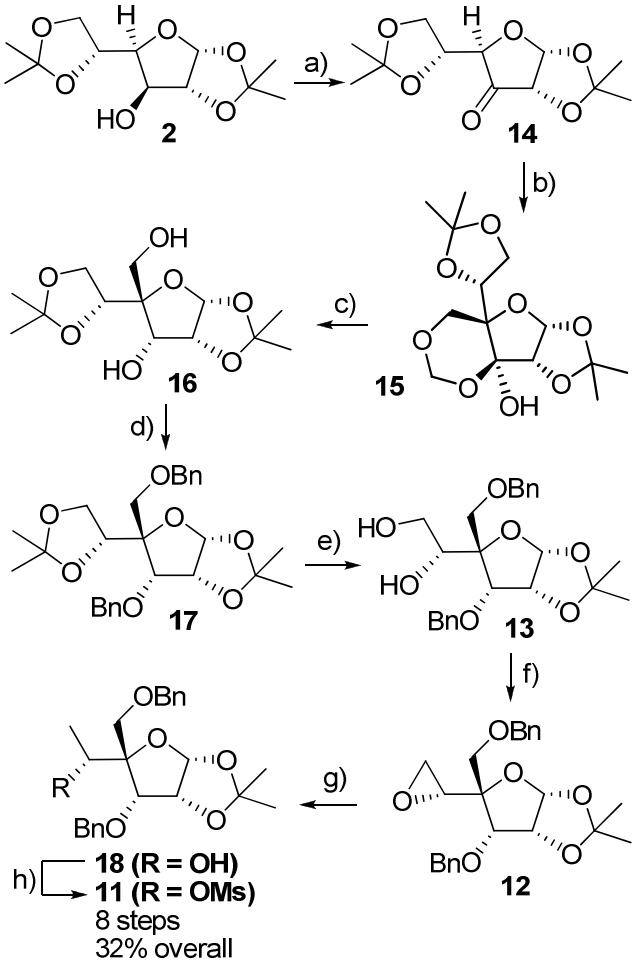


The epoxide **12** was expected to be obtainable from diol **13** which would require pre-installation of the 4'-quaternary stereocentre using aldol chemistry with formaldehyde as an electrophile and ketone **14** (obtainable in one step from **2**) as the enol nucleophile.

With the retrosynthetic plan in place, we began the forward synthesis with the oxidation of diacetone glucose **2** to ketone **14** using a modified version of a literature procedure as shown in **SCHEME 5**.[17] We were expecting the aldol reaction of **14** to produce regio- and stereochemical mixtures and only after numerous screening experiments did we find that ketone **14** undergoes an unusually selective reaction with aqueous formaldehyde in the presence of triethylamine using 2-methyltetrahydrofuran as solvent. This process, a variant of which was reported during the preparation of this manuscript by Kuwahara et al,[18] gives the acetal **15** in 65% yield after crystallisation. Structural assignment of **15** was not straightforward and single crystal x-ray analysis was used to confirm the relative stereochemistry. We were very interested to see that Kuwahara et al obtained mixtures of **15** (identified using 1D/2D-NMR) and the expected aldol,[18] suggesting that subtle reagent and solubility effects may be responsible for determining the outcome of reaction. Acetal **15** did not require any specific deprotection and underwent reduction with sodium borohydride to give diol **16** as a single stereoisomer in 87% yield. With safety and

scalability in mind, double benzyl protection of diol **16** was carried out using a phase transfer catalytic system to give acetonide **17** in 94% yield.

SCHEME 5. Synthesis of building block 9.



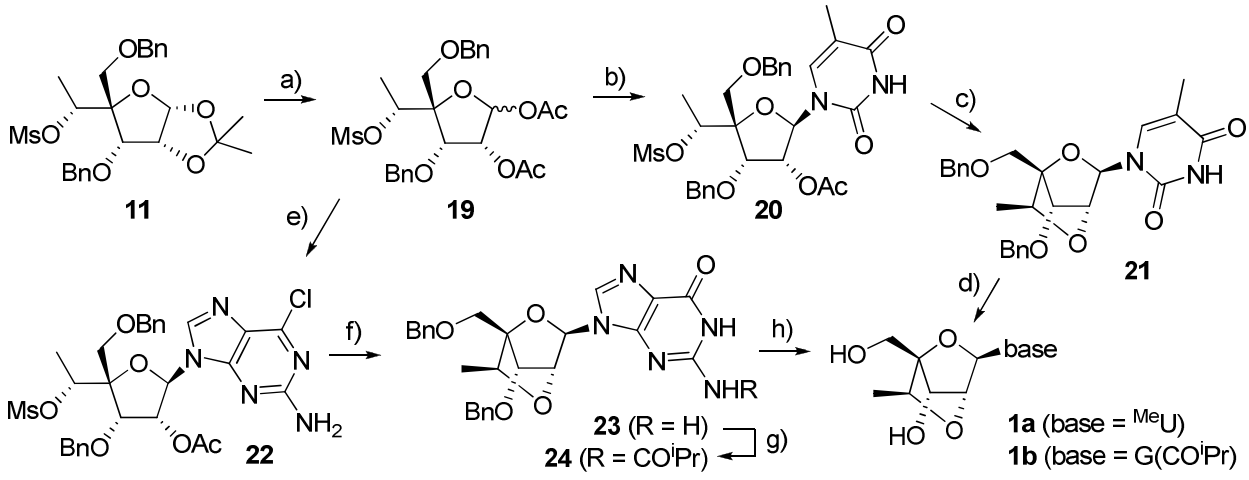
a) NaOCl (aq), TEMPO, KBr, DCM, 30 °C, 1 h, 81%; b) CH_2O (aq), Et_3N, 2-methyltetrahydrofuran, rt, 16 h, 65% after recrystallisation; c) $NaBH_4$, MeOH, 0 °C, 87%; d) BnBr, $[Bu_4N^+][HSO_4^-]$, NaOH (aq), 2-methyltetrahydrofuran, rt, 16 h, 94%; e) HCOOH, AcOH, H_2O, rt, 30 min, 94%; f) TsCl, $[PhCH_2(Et)_3N^+][Cl^-]$, NaOH (aq), toluene, rt, 3 h, 93%; g) $LiAlH_4$, THF, -5°C – rt, 2.5 h, 92%; h) MsCl, DMAP, MTBE, 0 °C, 1 h, 91%.

A number of approaches were attempted to achieve the selective removal of the primary acetonide in **17**. A procedure was developed using an aqueous mixture of formic and acetic acids which allowed diol **13** to be isolated in good yield and purity after work-up. We found that conversion of **13** to epoxide **12** could be

achieved using a modified 1-pot tosylation/substitution procedure giving **12** in 93% yield.[19] Full characterisation (including single crystal x-ray analysis) verified the structure and stereochemistry of **12**. Hydride reduction at the epoxide terminus and mesylation of the resulting secondary alcohol proceeded smoothly giving **11** in 84% yield over 2 steps.

With access to gram quantities of building block **11**, synthesis of the desired cEt nucleosides **1a** and **1b** was demonstrated using the two divergent sequences shown in **SCHEME 6**.

SCHEME 6. Synthesis of nucleosides 1a and 1b.



a) Ac_2O, H_2SO_4, EtOAc, rt, 24 h, quantitative; b) thymine, *N*,*O*-bis(trimethylsilyl)acetamide, TMS-OTf, toluene, 80 °C, 2 h, 82%; c) NaOH (s), MeOH, 40 °C, 1 h, 80%; d) H_2 (60 psig), $Pd(OH)_2/C$, EtOH, 25 °C, 6 h, 92%; e) 2-amino-6-chloropurine, *N*,*O*-bis(trimethylsilyl)acetamide, TMSOTf, toluene, 80 °C, 2 h, 64%; f) NaO^tBu (s), 3-hydroxypropionitrile, THF, 0 °C to rt, 16 h, 76%; g) isobutyric anhydride, Et_3N, DMAP, toluene, 105 °C, 16 h, 96% h) $HCOONH_4$, $Pd(OH)_2/C$, isopropanol, 80 °C, 90 min, 82%.

Concomitant acetonide cleavage and acetylation gave **19** as a ~1:2 mixture of *cis*- and *trans*- isomers. Vorbrüggen glycosidation using typical conditions gave **20** and **22** with excellent selectivity at the 1'-position (>95:5 anomeric ratio) by HPLC.[20] Solvolysis of the 2'-OAc groups in **20** and **22** and subsequent cyclisation proved straightforward using typical conditions.[2,6] Following amide formation for the purine series (**23**→**24**), the final proof of the synthetic approach was demonstrated successfully. Conversion of

21→1a was achieved using $Pd(OH)_2$ supported on carbon (Evonik E101) under hydrogen at 60 psig.[15] The same conditions did not give complete conversion of **24→1b** and a transfer hydrogenation using the same catalyst and ammonium formate in isopropanol was used.[21] From mesylate **11**, pyrimidine **1a** was generated in 60% yield over four steps; the purine analog **1b** (which comprises an extra amide forming step) was generated in 38% yield over five steps.

In summary, by comparison to the existing synthetic route to **1a** and **1b**, the approach above offers: i) a step count reduction from 23 to 17;[22] ii) increased yield to both **1a** (19% versus 8%) and **1b** (12% versus 6%); iii) greater atom efficiency with building block **11** (493 Da) having 29% lower molecular mass than **4** (691 Da) owing to the replacement of the protecting groups and iv) an improved purity profile with the avoidance of the related cMe analogs **6**. We are confident that this approach constitutes a general method for the synthesis of pyrimidine and purine cEt nucleosides and will ultimately help provide new oligonucleotide medicines to patients in the near future.

Experimental section

General Procedure. All of the processes described were carried out under nitrogen atmospheres without the requirement for oven- or flame-dried glassware. All solvents and reagents were used as supplied without any further treatment. Column chromatography was carried using automated systems with pre-packed silica columns. NMR spectra were recorded at 400 or 500 MHz and calibrated using residual undeuterated solvent as an internal reference: $CHCl_3$, δ 7.24 ppm; DMSO, δ 2.49 ppm; MeOH, δ 3.30 ppm and reported in parts per million relative to trimethylsilane at δ 0.00 ppm. NMR data is reported in the following format: chemical shift (multiplicity, coupling constant (Hz), integration) using the following abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, spt = septet, m = multiplet, br = broad, dd = doublet of doublets, dt = doublet of triplets, ddt = doublet of doublet of triplets. High resolution mass spectra (HRMS) were recorded on a quadrupole-TOF mass spectrometer using positive electrospray ionization.

(3a*R*,5*R*,6a*S*)-5-[(4*R*)-2,2-Dimethyl-1,3-dioxolan-4-yl]-2,2-dimethyl-3a,6a-dihydrofuro[2,3-d][1,3]dioxol-6-one (14)

To a solution of diacetone glucose **2** (322 g, 1.24 mol) in DCM (1.60 L) was added KBr (1.48 g, 124 mmol) and TEMPO (4.83 g, 30.9 mmol). The resulting mixture was heated to 30 °C with stirring and NaOCl (~15% aq, 768 mL, 1.55 mol) was added over ~1 h maintaining a temperature of 25 – 35 °C. *Note that additions carried out at lower temperatures can lead to accumulation of reactive mixtures and the potential for thermal runaway.* The layers were separated and the organic layer was washed sequentially with a solution of KI (12.9 g, 77.6 mmol) in HCl (0.50 M aq, 805 mL), $Na_2S_2O_3$ (aq, sat, 805 mL) and $NaHCO_3$ (~5% w/w aq, 805 mL). The combined organics were dried over $MgSO_4$ and concentrated under reduced pressure. The residue was azeodried three times by evaporation from toluene (500 mL) to give ketone **14** (261 g, 81%) as a brown liquid: ^{1}H NMR (400MHz, $CDCl_3$) δ 6.15 (d, J = 4.5 Hz, 1H), 4.42 - 4.33 (m, 3H), 4.07 - 4.01 (m, 2H), 1.47 (s, 3H), 1.45 - 1.43 (m, 3H), 1.35 (s, 6H) consistent with literature values.[17]

(3a*S*,3b*R*,7a*S*,8a*R*)-7a-[(4*R*)-2,2-Dimethyl-1,3-dioxolan-4-yl]-2,2-dimethyltetrahydro-3b*H*-[1,3]dioxolo[4,5]furo[3,2-*d*][1,3]dioxin-3b-ol (15)

To a solution of ketone **14** (39.8 g, 154 mmol) in 2-methyltetrahydrofuran (800 mL) was added Et_3N (215 mL, 1.54 mol) and HCHO (37% w/w aq, 398 mL, 14.7 mol). The mixture was stirred at rt for 16 h. The mixture was adjusted to pH 4 using NH_4Cl (aq, sat). After addition of EtOAc (200 mL) and H_2O (200 mL) the layers were separated. The aqueous layer was extracted twice using EtOAc (200 mL). The combined organics were dried over $MgSO_4$ and concentrated under reduced pressure. The residue was dissolved into EtOAc (130 mL) at 70 °C and hexane (400 mL) was added drop-wise over 30 min. The resulting turbid solution was allowed to cool to rt and left to stir for 16 h. The solid was collected via filtration and the filter cake was washed twice with hexane (400 mL). The solid was dried in-vacuo at 40 °C for 16 h to give acetal **15** (32.0 g, 65%) as a light yellow crystalline solid: mp 102.0 – 104.5 °C; ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.99 (d, *J* = 4.3 Hz, 1H), 4.95 (d, *J* = 6.2 Hz, 1H), 4.85 (dd, *J* = 0.8, 6.0 Hz, 1H), 4.82 (t, *J* = 6.9 Hz, 1H), 4.31 (d, *J* = 4.3 Hz, 1H), 4.15 (d, *J* = 13.0 Hz, 1H), 4.07 (d, *J* = 6.7 Hz, 2H), 3.78 (d, *J* = 13.2 Hz, 1H), 3.68 (s, 1H), 1.62 (s, 3H), 1.43 (s, 3H), 1.35 (s, 3H), 1.33 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 113.1, 108.4, 103.9, 99.8, 86.7, 83.8, 82.2, 76.4, 65.8, 65.0, 26.1, 25.9, 25.8, 23.8; HRMS (ESI-TOF) calcd for $C_{14}H_{22}NaO_8$ $[M + Na]^+$ *m/z* = 341.1207, found 341.1205. A crystal of **15** with approximate dimensions of 0.20×0.30×0.80 mm was mounted and automatically centered on a benchtop crystallographic system. Intensity measurements were performed using monochromated (doubly curved silicon crystal) MoK_α radiation (0.71073 Å) from a sealed microfocus tube. Generator settings were 50 kV, 1 mA. Data collection temperature was -73 °C. APEX2 software was used for preliminary determination of the unit cell. Determination of integrated intensities and unit cell refinement were performed using SAINT. The integration of the data yielded a total of 6381 reflections to a maximum θ angle of 24.39° (0.86 Å resolution). The constants for the triclinic unit cell are a = 6.098(4) Å, b = 7.378(5) Å, c = 9.239(5) Å, α = 106.18(2)°, β = 100.896(19)°, γ = 96.42(2)°, V = 386.0(4) Å^3. These are based upon the refinement of the XYZ-centroids of reflections above 20 σ(I). The calculated minimum and maximum transmission coefficients (based on crystal size) are 0.9150 and 0.9780. The

average residual for symmetry equivalent reflections is R_{int} = 7.66% and R_σ = 10.28%. XPREP determined the space group to be P 1, with Z = 1 for the formula unit,$C_{14}H_{22}O_8$. The structure was solved with XS and subsequent structure refinements were performed with XL. The final anisotropic full-matrix least-squares refinement on F_o^2 with 204 variables converged at R_1 = 6.07% for the observed data and wR_2 = 16.68% for all data. The goodness-of-fit was 0.971. The largest peak on the final difference electron density synthesis was 0.23 $e^-/Å^3$ and the deepest hole was -0.34 $e^-/Å^3$ with an RMS deviation of 0.06 $e^-/Å^3$. On the basis of the final model, the calculated density is 1.369 g/cm^3 and F(000) = 170. Absolute stereochemistry was assigned relative to the (1'*R*,2'*R*) stereocentres derived from the parent structure diacetone glucose **2**.

(3a*R*,5*R*,6*S*,6a*R*)-5-[(4*R*)-2,2-Dimethyl-1,3-dioxolan-4-yl]-5-(hydroxymethyl)-2,2-dimethyl-6,6a-dihydro-3a*H*-furo[2,3-d][1,3]dioxol-6-ol (16)

A suspension of acetal **15** (204 g, 640 mmol) in MeOH (2.90 L) was cooled to 0 °C. Maintaining a temperature below 8 °C, $NaBH_4$ (22.9g, 605 mmol) was added portion-wise over 65 min. The resulting solution was stirred for 30 min at 0 °C, and the reaction was then quenched by careful addition of H_2O (20.0 mL). Further H_2O (2.70 L) was added and the resulting solution extracted four times with DCM (2.50 L). The combined organics were dried over $MgSO_4$ and concentrated under reduced pressure to give diol **16** (171 g, 87%) as an off-white solid: ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.92 (d, *J* = 4.1 Hz, 1H), 4.73 (dd, *J* = 4.1, 6.4 Hz, 1H), 4.59 (dd, *J* = 6.5, 7.5 Hz, 1H), 4.33 (t, *J* = 6.5 Hz, 1H), 4.16 (dd, *J* = 7.5, 9.3 Hz, 1H), 3.91 (dd, *J* = 6.5, 9.3 Hz, 1H), 3.79 (d, *J* = 11.9 Hz, 1H), 3.61 (d, *J* = 11.5 Hz, 1H), 2.80 (d, *J* = 6.8 Hz, 1H), 2.21 - 2.05 (m, 1H), 1.75 - 1.66 (m, 1H), 1.63 (s, 3H), 1.46 (s, 3H), 1.41 (s, 3H), 1.35 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 114.2, 108.9, 105.6, 90.5, 80.4, 77.5, 71.6, 65.8, 62.7, 27.0, 26.9, 26.0, 24.1; HRMS (ESI-TOF) calcd for $C_{13}H_{22}NaO_7$ $[M + Na]^+$ *m/z* = 313.1258, found 313.1257.

(3a*R*,5*S*,6*S*,6a*R*)-6-Benzyloxy-5-(benzyloxymethyl)-5-[(4*R*)-2,2-dimethyl-1,3-dioxolan-4-yl]-2,2-dimethyl-6,6a-dihydro-3a*H*-furo[2,3-d][1,3]dioxole (17)

Diol **16** (1.03 g, 3.55 mmol) and tetrabutylammonium hydrogensulfate (0.24 g, 0.71 mmol) were dissolved in 2-methyltetrahydrofuran (10.3 mL). Following addition of NaOH (~50% w/w aq, 1.86 mL,

35.2 mmol) BnBr (0.86 mL, 7.22 mmol) was added drop-wise over 30 min. The reaction mixture was stirred rapidly for 16 h. After addition of H_2O (5.10 mL), the phases were separated. The aqueous phase was extracted twice with MTBE (10.3 mL) and the combined organics were washed with NaCl (aq, sat, 5.10 mL), dried over $MgSO_4$ and concentrated under reduced pressure to give acetonide **17** (1.56g, 94%) as a clear oil: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 - 7.26 (m, 10H), 5.85 (d, *J* = 3.8 Hz, 1H), 4.80 (dd, *J* = 6.5, 7.5 Hz, 1H), 4.77 (d, *J* = 11.5 Hz, 1H), 4.69 (dd, *J* = 3.9, 5.2 Hz, 1H), 4.55 (d, *J* = 11.8 Hz, 1H), 4.47 (d, *J* = 11.9 Hz, 1H), 4.43 (d, *J* = 11.5 Hz, 1H), 4.27 (d, *J* = 5.3 Hz, 1H), 4.07 (dd, *J* = 7.6, 9.2 Hz, 1H), 3.79 (dd, *J* = 6.4, 9.2 Hz, 1H), 3.69 (d, *J* = 10.5 Hz, 1H), 3.59 (d, *J* = 10.5 Hz, 1H), 1.62 (s, 3H), 1.41 (s, 3H), 1.39 (s, 3H), 1.31 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 138.1, 137.5, 128.5, 128.4, 128.1, 128.0, 127.7, 127.6, 114.1, 108.3, 105.4, 88.8, 80.4, 78.5, 77.4, 73.9, 72.9, 70.1, 66.1, 27.2, 27.1, 26.0, 23.9; HRMS (ESI-TOF) calcd for $C_{27}H_{34}NaO_7$ $[M + Na]^+$ *m/z* = 493.2197, found 493.2196.

3-*O*-Benzyl-4-[(benzyloxy)methyl]-1,2-*O*-(1-methylethylidene)-α-D-gulofuranose (13)

Acetonide **17** (2.00 g, 4.25 mmol) was dissolved in a pre-made mixture of AcOH (11.6 mL, 203 mmol), HCOOH (4.80 mL, 127 mmol) and H_2O (3.60 mL, 200 mmol) and the mixture was stirred at rt for 30 min. The mixture was poured on to ice-cold NaOH (~50% w/w aq, 26.9 mL, 350 mmol) and stirred at 0 °C for 10 min. After addition of DCM (100 mL) and H_2O (100 mL) the layers were separated. The aqueous layer was extracted twice with DCM (100 mL). The combined organics were dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by column chromatography on silica gel (30% - 50% EtOAc in hexane) to give diol **13** (1.72 g, 94%) as a colourless oil: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.36 - 7.24 (m, 10H), 5.81 (d, *J* = 3.8 Hz, 1H), 4.81 (d, *J* = 11.8 Hz, 1H), 4.67 (dd, *J* = 3.9, 5.3 Hz, 1H), 4.50 (d, *J* = 11.6 Hz, 2H), 4.42 (d, *J* = 11.9 Hz, 1H), 4.34 (d, *J* = 5.4 Hz, 1H), 4.23 (ddd, *J* = 2.5, 3.9, 6.4 Hz, 1H), 3.78 (ddd, *J* = 3.9, 7.9, 11.3 Hz, 1H), 3.69 (ddd, *J* = 4.4, 6.2, 10.9 Hz, 1H), 3.48 (q, *J* = 10.2 Hz, 2H), 3.22 (d, *J* = 2.3 Hz, 1H), 2.43 (dd, *J* = 4.5, 7.9 Hz, 1H), 1.66 (s, 3H), 1.37 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 137.7, 137.1, 128.6, 128.5, 128.2, 127.9, 127.8, 127.7, 114.3, 104.8, 88.4, 79.1, 79.1, 73.7, 72.9, 71.5, 71.2, 62.7, 26.7, 26.5; HRMS (ESI-TOF) calcd for $C_{24}H_{30}NaO_7$ $[M + Na]^+$ *m/z* = 453.1884, found 453.1880.

5,6-Anhydro-3-*O*-benzyl-4-[(benzyloxy)methyl]-1,2-*O*-(1-methylethylidene)-α-D-gulofuranose (12)

To a solution of diol **13** (5.99 g, 13.9 mmol) and benzyltriethylammonium chloride (317 mg, 1.39 mmol) in toluene (120 mL) was added NaOH (~50% w/w aq, 47.9 mL, 908 mmol). A solution of TsCl (2.84 g, 14.7 mmol) in toluene (120 mL) was added drop-wise over a period of ~2 h. The mixture was stirred for a further 1 h before H_2O (120 mL) was added and the layers were separated. The aqueous layer was extracted twice with toluene (120 mL). The combined organics were washed with H_2O (60.0 mL), NaCl (aq, sat, 60.0 mL), dried over $MgSO_4$ and concentrated under reduced pressure to give epoxide **12** (5.36 g, 93%) as a colourless crystalline solid: mp 72.0 – 72.5 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.34 - 7.22 (m, 10H), 5.76 (d, *J* = 3.7 Hz, 1H), 4.79 (d, *J* = 12.3 Hz, 1H), 4.63 (dd, *J* = 3.8, 5.0 Hz, 1H), 4.56 (d, *J* = 12.2 Hz, 1H), 4.49 (d, *J* = 12.0 Hz, 1H), 4.41 (d, *J* = 12.0 Hz, 1H), 4.31 (d, *J* = 5.0 Hz, 1H), 3.53 (dd, *J* = 2.9, 4.2 Hz, 1H), 3.38 (d, *J* = 10.5 Hz, 1H), 3.24 (d, *J* = 10.5 Hz, 1H), 2.78 - 2.73 (m, 2H), 1.67 (s, 3H), 1.36 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 137.9, 137.8, 128.4, 127.9, 127.8, 127.7, 127.6, 114.0, 104.5, 85.8, 79.5, 78.7, 73.7, 72.6, 69.8, 53.0, 44.6, 26.9, 26.5; HRMS (ESI-TOF) calcd for $C_{24}H_{28}NaO_6$ $[M + Na]^+$ *m/z* = 435.1778, found 435.1781.

A crystal of **12** with approximate dimensions of 0.20×0.60×1.00 mm was mounted and automatically centred on a bench-top crystallographic system. Intensity measurements were performed using monochromated (doubly curved silicon crystal) MoK_α radiation (0.71073 Å) from a sealed microfocus tube. Generator settings were 50 kV, 1 mA. Data collection temperature was -73 °C. APEX2 software was used for preliminary determination of the unit cell. Determination of integrated intensities and unit cell refinement were performed using SAINT. The integration of the data yielded a total of 12079 reflections to a maximum θ angle of 25.07° (0.84 Å resolution). The constants for the monoclinic unit cell are a = 10.6029(15) Å, b = 8.4121(11) Å, c = 13.0245(17) Å, β = 109.451(5)°, V = 1095.4(3) Å^3.They are based upon the refinement of the XYZ-centroids of 3851 reflections above 20.0 I/σ(I) with 2.93° ≤ θ ≤23.20°. Data were corrected for absorption effects with SADABS using the multi-scan technique. The ratio of minimum to maximum apparent transmission is 64.6:100.The average residual for symmetry equivalent reflections is R_{int} = 5.24% and R_σ = 4.84%. XPREP determined the space group to be P 1 21 1,

with Z = 2 for the formula unit, $C_{24}H_{28}O_6$. The structure was solved with SHELXTL XT and subsequent structure refinements were performed with XShell. The final anisotropic full-matrix least-squares refinement on F_o^2 with 273 variables converged at R_1 = 3.79% for the observed data and wR_2 = 12.85% for all data. The goodness-of-fit was 0.847. The largest peak on the final difference electron density synthesis was 0.30 $e^-/Å^3$ and the deepest hole was-0.26 $e^-/Å^3$ with an RMS deviation of 0.09 $e^-/Å^3$. On the basis of the final model, the calculated density is 1.251 g/cm^3 and F(000) = 440. Absolute stereochemistry was assigned relative to the (1'*R*,2'*R*) stereocentres derived from the parent structure diacetone glucose **2**.

3-*O*-Benzyl-4-[(benzyloxy)methyl]-6-deoxy-1,2-*O*-(1-methylethylidene)-α-D-gulofuranose (18)

Epoxide **12** (3.56 g, 8.63 mmol) was dissolved in THF (66.8 mL, 820 mmol) and the solution was cooled to -5 °C. Drop-wise addition of $LiAlH_4$ solution (~1 M in THF, 6.88 mL, 6.88 mmol) was carried out over 15 min and the mixture was stirred at -5 °C for a further 10 min. The mixture was allowed to warm to rt and was stirred for 2 h. The mixture was cooled to 0 °C and H_2O (0.26 mL) was added then the mixture was stirred for a further 5 min. After the addition of NaOH (~15% w/w aq, 0.26 mL), H_2O (0.79 mL) was added and the resulting solution was allowed to warm to rt and stirred for 2 h. After addition of $MgSO_4$ (3.56 g, 29.6 mmol) the mixture was stirred for 30 min. The resulting suspension was filtered through a pad of Celite, and the filter cake was washed with EtOAc (20.0 mL). The filtrate was concentrated under reduced pressure to give alcohol **18** (3.17 g, 92%) as a colourless solid: ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 - 7.24 (m, 10H), 5.82 (d, *J* = 3.9 Hz, 1H), 4.80 (d, *J* = 11.9 Hz, 1H), 4.68 (dd, *J* = 4.0, 5.4 Hz, 1H), 4.53 (d, *J* = 6.7 Hz, 1H), 4.50 (d, *J* = 6.6 Hz, 1H), 4.42 (d, *J* = 12.0 Hz, 1H), 4.37 - 4.30 (m, 1H), 3.54 (d, *J* = 10.1 Hz, 1H), 3.45 (d, *J* = 10.0 Hz, 1H), 3.08 - 3.00 (m, 1H), 1.64 (s, 3H), 1.37 (s, 3H), 1.20 (d, *J* = 6.5 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 138.0, 137.6, 128.5, 128.4, 128.0, 127.8, 127.7, 127.6, 114.1, 104.5, 89.2, 79.5, 78.8, 73.7, 72.8, 70.6, 67.2, 26.8, 26.6, 17.1; HRMS (ESI-TOF) calcd for $C_{24}H_{30}NaO_6$ $[M + Na]^+$ *m/z* = 437.1935, found 437.1932.

3-*O*-Benzyl-4-[(benzyloxy)methyl]-6-deoxy-1,2-*O*-(1-methylethylidene)-5-*O*-(methylsulfonyl)-α-D-gulofuranose (11)

To a solution of alcohol **18** (1.14 g, 2.75 mmol) and DMAP (52.0 mg, 0.41 mmol) in MTBE (6.00 mL) was added Et_3N (0.77 mL, 5.50 mmol) and the resulting solution was cooled to 0 °C. Methanesulfonyl chloride (0.32 mL, 4.13 mmol) was added drop-wise over 5 min and the reaction was stirred at 0 °C for 1 h. After addition of MTBE (1.25 mL) and H_2O (1.25 mL), the layers were separated. The aqueous layer was extracted twice with MTBE (5.00 mL). The combined organics were washed twice with H_2O (5.00 mL), NaCl (aq, sat, 5 mL), dried over magnesium sulfate and concentrated under reduced pressure. The residue was purified by column chromatography on silica gel (20% - 30% EtOAc in hexane) to give mesylate **11** (1.24 g, 91%) as a colourless oil: ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 - 7.22 (m, 10H), 5.87 (d, *J* = 4.3 Hz, 1H), 5.25 (q, *J* = 6.6 Hz, 1H), 4.84 - 4.74 (m, 2H), 4.55 (d, *J* = 11.9 Hz, 1H), 4.45 - 4.42 (m, 2H), 4.20 (d, *J* = 5.5 Hz, 1H), 3.68 (d, *J* = 10.1 Hz, 1H), 3.61 (d, *J* = 10.1 Hz, 1H), 3.14 (s, 3H), 1.63 (s, 3H), 1.46 (d, *J* = 6.6 Hz, 3H), 1.38 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 137.6, 137.1, 128.5, 128.5, 128.1, 128.0, 127.7, 114.3, 105.4, 88.3, 82.6, 80.6, 78.3, 73.9, 73.1, 69.7, 38.7, 27.0, 26.5, 18.8; HRMS (ESI-TOF) calcd for $C_{25}H_{32}NaO_8S$ $[M + Na]^+$ *m/z* = 515.1710, found 515.1705.

1,2-Di-*O*-acetyl-3-*O*-benzyl-4-[(benzyloxy)methyl]-6-deoxy-5-*O*-(methylsulfonyl)-D-gulofuranose (19)

To a solution of mesylate **11** (1.00 g, 2.03 mmol) in EtOAc (4.00 mL) was added Ac_2O (576 µL, 6.09 mmol) and the solution was cooled to 0 °C. After addition of H_2SO_4 (~98%, 22.2 µL, 0.406 mmol) the reaction mixture was allowed to warm to rt and stirred for 24 h. After addition of $NaHCO_3$ (aq, sat, 2.00 mL) the mixture was stirred at rt for 10 min. The layers were separated and the aqueous layer was extracted three times with EtOAc (5.00 mL). The combined organics were dried over $MgSO_4$ and was concentrated under reduced pressure to give acetate **19** (1.09 g, quantitative) as a colourless oil. Compound isolated as a 2:1 mixture of 1'-epimers. A sample was purified by column chromatography on silica gel (10% - 50% EtOAc in hexane) to give samples of both isomers which were assigned based on coupling constants, *trans*-**19** ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 - 7.23 (m, 10H), 6.15 (d, *J* = 1.4 Hz, 1H), 5.46 - 5.38 (m, 1H), 5.17 (q, *J* = 6.6 Hz) 1.4H), 4.60 (d, *J* = 11.2 Hz, 1H), 4.55 - 4.45 (m, 4H), 3.68 (d, *J* = 10.2 Hz, 1H), 3.53 (d, *J* = 10.2 Hz, 1H), 3.02 (s, 3H), 2.14 (s, 3H), 1.45 (d, *J* = 6.6 Hz, 3H); ^{13}C

NMR (125 MHz, $CDCl_3$) δ 169.8, 169.3, 137.6, 136.8, 128.6, 128.5, 128.3, 128.1, 127.9, 127.7, 97.5, 88.8, 81.0, 78.6, 74.6, 74.2, 73.6, 69.4, 38.5, 20.9, 20.8, 18.0; *cis*-**19** ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.38 - 7.27 (m, 10H), 6.44 (d, *J* = 4.8 Hz, 1H), 5.46 - 5.38 (m, 1H), 5.25 - 5.15 (m, 1H), 4.73 (d, *J* = 11.5 Hz, 1H), 4.61 (m, 1H), 4.55 - 4.45 (m, 3H), 4.30 (d, *J* = 5.3 Hz, 1H), 3.72 (d, *J* = 10.3 Hz, 1H), 3.57 (d, *J* = 10.3 Hz, 0.4H), 3.03 (s, 1.2H), 3.02 (s, 3H), 2.14 (s, 3H), 1.99 (s, 3H), 1.34 (d, *J* = 6.6 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.5, 169.9, 137.6, 137.3, 128.6, 128.4, 128.0, 127.8, 127.7, 127.1, 94.2, 89.5, 81.8, 78.1, 74.5, 73.9, 73.2, 69.3, 38.4, 21.2, 20.5, 18.4; HRMS (ESI-TOF) calcd for $C_{26}H_{32}NaO_{10}S$ $[M + Na]^+$ *m/z* = 559.1608, found 559.1600.

1-{2-*O*-Acetyl-3-*O*-benzyl-4-[(benzyloxy)methyl]-6-deoxy-5-*O*-(methylsulfonyl)-β-D-gulofuranosyl}-5-methylpyrimidine-2,4(1*H*,3*H*)-dione (20)

To a solution of acetate **19** (1.00 g, 1.86 mmol) in MeCN (5.00 mL) were added thymine (285 mg, 2.24 mmol) and *N*,*O*-bis(trimethylsilyl)acetamide (1.38 mL, 5.59 mmol) and the resulting mixture was heated at 40 °C for 1 h. The mixture was cooled to rt and TMSOTf (452 μL, 2.42 mmol) was added before the mixture was heated at 80 °C for 1 h. The reaction mixture was cooled to 0 °C and the pH was adjusted to pH 6-7 using NaOH (15% w/w aq). After addition of EtOAc (10.0 mL) and H_2O (10.0 mL) the layers were separated. The aqueous layer was extracted twice with EtOAc (10.0 mL). The combined organics were washed with NaCl (aq, sat, 10.0 mL), dried over $MgSO_4$ and concentrated under reduced pressure to give pyrimidine **20** (0.92 g, 82%) as a colourless foam: ^{1}H NMR (500MHz, $CDCl_3$) δ 7.44 - 7.28 (m, 12H), 6.23 (d, *J* = 6.6 Hz, 1H), 5.64 (t, *J* = 6.1 Hz, 1H), 5.07 (q, *J* = 6.4 Hz, 1H), 4.68 (d, *J* = 10.9 Hz, 1H), 4.60 (s, 2H), 4.50 (d, *J* = 10.9 Hz, 1H), 4.46 (d, *J* = 5.7 Hz, 1H), 3.79 (d, *J* = 9.9 Hz, 1H), 3.65 (d, *J* = 9.9 Hz, 1H), 2.94 (s, 3H), 2.09 (s, 3H), 1.57 (s, 3H), 1.38 (d, *J* = 6.5 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.4, 136.8, 136.6, 136.1, 128.8, 128.7, 128.5, 128.4, 128.2, 127.8, 124.1, 111.7, 87.8, 87.0, 80.1, 79.1, 75.1, 74.5, 74.1, 70.2, 38.22, 20.8, 17.4, 12.1.

1-[(1*R*,3*S*,4*S*,6*R*,7*S*)-7-Benzyloxy-4-(benzyloxymethyl)-3-methyl-2,5-dioxabicyclo[2.2.1]heptan-6-yl]-5-methyl-pyrimidine-2,4-dione (21)

To a solution of pyrimidine **20** (1.10 g, 1.83 mmol) in MeOH (4.40 mL) was added NaOH (220 mg, 5.50 mmol) and the suspension was heated at 40 °C for 1 h. The volatiles were removed under reduced pressure and the residue was taken up in EtOAc (10 ml) and H_2O (10.0 mL). The layers were separated and the aqueous layer extracted twice with EtOAc (10.0 mL). The combined organics were washed with NaCl (aq, sat, 10.0 mL), dried over $MgSO_4$ and concentrated under reduced pressure to give pyrimidine **21** (678 mg, 80%) as a light yellow oil: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.29 (s, 1H), 7.45 (d, *J* = 1.0 Hz, 1H), 7.41 - 7.28 (m, 10H), 5.61 (s, 1H), 4.64 (m, 3H), 4.56 (s, 1H), 4.50 (d, *J* = 11.6 Hz, 1H), 4.13 (q, *J* = 6.7 Hz, 1H), 3.90 - 3.87 (m, 3H), 1.65 (d, *J* = 0.7 Hz, 3H), 1.32 (d, *J* = 6.7 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 163.5, 150.0, 137.5, 136.8, 134.9, 128.6, 128.5, 128.1, 128.1, 127.9, 127.7, 110.1, 88.0, 87.0, 81.2, 73.9, 72.4, 64.9, 16.4, 12.4; HRMS (ESI-TOF) calcd for $C_{26}H_{29}N_2O_6$ $[M + H]^+$ *m/z* = 465.2020, found 465.2022.

1-[(1*R*,3*S*,4*R*,6*R*,7*S*)-7-Hydroxy-4-(hydroxymethyl)-3-methyl-2,5-dioxabicyclo[2.2.1]heptan-6-yl]-5-methyl-pyrimidine-2,4-dione (1a)

To a solution of pyrimidine **21** (100 mg, 215 μmol) in EtOH (5.00 mL) was added $Pd(OH)_2$ on carbon (Evonik type E101; 20% w/w loading, 27.5 mg) and the mixture was hydrogenated at 25 °C and 60 psig for 6 h. The reaction mixture was filtered through a plug of Celite, washing through with EtOH (10.0 mL) and the filtrate was concentrated under reduced pressure. The residue was purified by column chromatography on silica gel (0 - 10% MeOH in DCM) to give nucleoside **1a** (56.0 mg, 92%) as a colourless solid. All data in full agreement with authentic pharmaceutical samples: ^{1}H NMR (400MHz, MeOH-d_4) δ 7.70 (s, 1H), 5.50 (s, 1H), 4.34 (s, 1H), 4.11 - 4.01 (m, 2H), 3.96 (s, 2H), 1.88 (s, 3H), 1.32 (d, *J* = 6.7 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.9, 150.0, 134.9, 108.3. 89.0. 86.0. 80.3. 79.3. 69.6, 56.1, 16.3, 12.3; HRMS (ESI-TOF) calcd for $C_{12}H_{17}N_2O_6$ $[M + H]^+$ *m/z* = 285.1081, found 285.1087.

[(2*R*,3*R*,4*S*,5*R*)-2-(2-Amino-6-chloro-purin-9-yl)-4-benzyloxy-5-(benzyloxymethyl)-5-[(1*R*)-1-methylsulfonyloxyethyl]tetrahydrofuran-3-yl] acetate (22)

To a suspension of acetate **19** (1.08 g, 2.01 mmol) in toluene (8.64 mL) were added 4-amino-6-chloropurine (375 mg, 2.21 mmol) and *N*,*O*-bis(trimethylsilyl)acetamide (1.48 mL, 6.04 mmol) and the mixture was heated at 80 °C for 1 h. The mixture was cooled to rt and TMSOTf (729 µL, 4.03 mmol) was added drop-wise over 5 min and the mixture was heated to 80 °C for 70 min. The mixture was cooled to 0 °C and the pH was adjusted to pH 6-7 using NaOH (~15% w/w aq). After the addition of EtOAc (10.0 mL) and H_2O (10.0 mL) the layers were separated. The aqueous layer was extracted twice with EtOAc (10.0 mL). The combined organics were washed with NaCl (aq, sat, 10.0 mL), dried over $MgSO_4$ and concentrated under reduced pressure to give purine **22** (0.83 g, 64%) as a colourless foam: ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.92 (s, 1H), 7.43 - 7.29 (m, 10H), 6.22 (d, *J* = 6.3 Hz, 1H), 6.02 (t, *J* = 6.0 Hz, 1H), 5.22 - 5.19 (m, 3H), 4.66 (d, *J* = 10.9 Hz, 1H), 4.62 (d, *J* = 5.7 Hz, 1H), 4.59 - 4.53 (m, 3H), 3.71 (d, *J* = 10.0 Hz, 1H), 3.64 (d, *J* = 10.0 Hz, 1H), 2.90 (s, 3H), 2.03 (s, 3H), 1.41 (d, *J* = 6.6 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 171.2, 170.0, 159.2, 153.6, 151.5, 141.0, 136.9, 136.6, 128.8, 128.7, 128.5, 128.3, 128.2, 127.9, 125.4, 88.2, 85.4, 80.3, 79.1, 75.0, 74.8, 73.9, 69.7, 60.4, 38.5, 21.1, 20.7, 17.4; HRMS (ESI-TOF) calcd for $C_{29}H_{33}ClN_5O_8S$ $[M + H]^+$ *m/z* = 646.1733, found 646.1730.

2-Amino-9-[(1*R*,3*S*,4*S*,6*R*,7*S*)-7-benzyloxy-4-(benzyloxymethyl)-3-methyl-2,5-dioxabicyclo[2.2.1]heptan-6-yl]-1*H*-purin-6-one (23)

A solution of NaO^tBu (700 mg, 7.26 mmol) in anhydrous THF (5.00 mL) was placed under an inert atmosphere and cooled to 0 °C. After stirring for 10 min, 3-hydroxypropionitrile (496 µL, 7.26 mmol) was added drop-wise and the resulting mixture was stirred at 0 °C for 30 min. A solution of purine **22** (816 mg, 1.26 mmol) in anhydrous THF (5.00 mL) was added drop-wise and the resulting mixture was stirred at 0 °C for 1 h before warming to rt and stirring for 16 h. The reaction mixture was cooled to 0 °C and H_2O (2.00 mL) was added drop-wise. After addition of EtOAc (10.0 mL), the layers were separated. The organic layer was washed with NH_4Cl (aq, sat, 5.00 mL), dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by column chromatography on silica gel (0% - 10% MeOH in DCM) to give aniline **23** (472 mg, 76%) as a yellow solid: ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.76 - 10.54 (m, 1H), 7.71 (s, 1H), 7.37 - 7.26 (m, 10H), 6.58 (br s, 2H), 5.73 (s, 1H), 4.72 (s, 1H), 4.67 - 4.57

(m, 4H), 4.28 (s, 1H), 4.13 (q, *J* = 6.6 Hz, 1H), 3.93 (d, *J* = 11.6 Hz, 1H), 3.88 (d, *J* = 11.5 Hz, 1H), 1.26 (d, *J* = 6.6 Hz, 3H); ^{13}C NMR (126 MHz, DMSO-d_6) δ 156.7, 153.9, 150.5, 138.0, 137.8, 134.1, 128.3, 128.3, 127.6, 127.6, 127.6, 127.4, 116.7, 87.0, 84.4, 80.5, 78.7, 77.0, 72.8, 71.4, 65.6, 16.3; HRMS (ESI-TOF) calcd for $C_{26}H_{28}N_5O_5$ $[M + H]^+$ *m/z* = 490.2085, found 490.2090.

_N_-[9-[(1_R_,3_S_,4_S_,6_R_,7_S_)-7-Benzyloxy-4-(benzyloxymethyl)-3-methyl-2,5-dioxabicyclo[2.2.1]heptan-6-yl]-6-oxo-1_H_-purin-2-yl]-2-methyl-propanamide (24)

To a solution of purine **23** (287 mg, 0.59 mmol) in toluene (2.00 mL) were added triethylamine (245 µL, 1.76 mmol) and DMAP (36.0 mg, 0.29 µmol). Isobutyric anhydride (291 µL, 1.76 mmol) was added and the reaction mixture was heated at 105 °C for 16 h. The mixture was cooled to 40 °C, MeOH (175 µL) was added and the resulting mixture was stirred at 40 °C for 30 min. The solution was cooled to rt, NH_4Cl (aq, sat, 5.00 mL) was added and the mixture was stirred for 20 min. The layers were separated and the aqueous layer was extracted twice with toluene (5.00 mL), dried over $MgSO_4$ and concentrated under reduced pressure to give amide **24** (316 mg, 96%) as a pale-brown solid: ^{1}H NMR (500 MHz, $CDCl_3$) δ 12.01 (s, 1H), 8.55 (br s, 1H), 7.76 (s, 1H), 7.37 - 7.20 (m, 10H), 5.76 (s, 1H), 4.62 (s, 2H), 4.55 – 4.44 (m, 3H), 4.21 (d, *J* = 6.7 Hz, 1H), 4.09 (s, 1H), 3.89 - 3.80 (m, 2H), 2.65 (quin, *J* = 6.9 Hz, 1H), 1.33 (d, *J* = 6.7 Hz, 3H), 1.27 (dd, *J* = 4.3, 6.9 Hz, 7H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 178.6, 155.5, 147.5, 147.0, 137.4, 136.8, 136.0, 128.6, 128.5, 128.1, 128.0, 127.7, 127.7, 121.8, 87.8, 85.9, 81.6, 78.3, 77.6, 73.9, 72.7, 65.1, 36.6, 19.0, 16.4; HRMS (ESI-TOF) calcd for $C_{30}H_{34}N_5O_6$ $[M + H]^+$ *m/z* = 560.2504, found 560.2497.

_N_-[9-[(1_R_,3_S_,4_R_,6_R_,7_S_)-7-Hydroxy-4-(hydroxymethyl)-3-methyl-2,5-dioxabicyclo[2.2.1]heptan-6-yl]-6-oxo-1_H_-purin-2-yl]-2-methyl-propanamide (1b)

To a solution of amide **24** (38.6 mg, 0.07 mmol) in ^{i}PrOH (4.00 mL) was added ammonium formate (250 mg, 3.97 mmol) and $Pd(OH)_2$ on carbon (Evonik type E101; 20% w/w loading, 25 mg). The mixture was sealed in a pressure tube and heated to 80 °C with stirring for 90 min. The mixture was cooled and the catalyst removed via filtration through celite which was rinsed with ^{i}PrOH (2.00 mL). The resulting solution was concentrated in-vacuo and purified by column chromatography on silica gel (5% - 10%

MeOH in DCM) to give nucleoside **1b** (21.3, 82%) as a white solid. All data in full agreement with authentic pharmaceutical samples: ^{1}H NMR (400 MHz, MeOH-d_4) δ 8.07 (s, 1H), 5.84 (s, 1H), 4.54 (s, 1H), 4.27 (s, 1H), 4.16 (q, *J* = 6.8 Hz, 1H), 4.00 (s, 2H), 2.71 (spt, *J* = 6.9 Hz, 1H), 1.37 (d, *J* = 6.8 Hz, 3H), 1.22 (d, *J* = 6.9 Hz, 6H); ^{13}C NMR (101 MHz, MeOH-d_4) δ 180.4, 156.0, 148.5, 148.0, 137.0, 120.2, 89.0, 86.0, 81.2, 80.1, 70.8, 56.8, 35.6, 17.9, 15.3; HRMS (ESI-TOF) calcd for $C_{16}H_{22}N_5O_6$ $[M + H]^+$ *m/z* = 380.1565, found 380.1561.

Supporting Information. NMR spectra (^{1}H and ^{13}C) are provided for all synthesised compounds. HPLC traces are provided for crude reaction mixtures of compounds **19**, **20** and **22** to illustrate selectivity. ORTEP diagrams and .cif files are provided for compounds **12** and **15**. This material is available free of charge via the Internet at http://pubs.acs.org.

Acknowledgements

The authors would like to thank Andy McPherson, Michael Migawa and Lijan Chen for useful discussions along with Anthony Bristow and Helen Wheatcroft for analytical assistance.

References and footnotes

[1] Tamm, I.; Dörken, B.; Hartmann, G. *Lancet* **2001**, *358*, 489 – 497.

[2] Seth, P. P.; Vasques, G.; Allerson, C. A.; Berdeja, A.; Gaus, H.; Kinberger, G. A.; Prakash, T. P.; Migawa, M. T.; Bhat, B.; Swayze, E. E. *J. Org. Chem.* **2010**, *75*, 1569 – 1581.

[3] Gibson, N. W. J. *R. Coll. Physicians Edinburgh* **2014**, *44*, 196.

[4] Hong, D.; Kim, Y.; Younes, A.; Nemunaitis, J.; Fowler, N.; Hsu, J.; Zhou, T.; Fayad, L.; Zhang, N.; Piha-Paul, S.; Woessner, R.; Nadella, M. V.; Lawson, D.; Reimer, C.; Jo, M.; Schmidt, J.; Xiao, X.; Greenlee, S.; Hung, G.; Yamashita, M.; Blakey, D. C.; Monia, B. P.; Macleod, A. R.; Kurzrock, R. *Cancer Res.* **2014**, *74*, LB-227.

[5] Betson, M.; Allanson, N.; Wainwright, P. *Org. Biomol. Chem.* **2014**, *12*, 9291 – 9306.

[6] Salinas, J. C.; Migawa, M. T.; Merner, B. L. Hanessian, S. *J. Org. Chem.* **2014**, *79*, 11651 – 11660.

[7] Seth, P. P.; Yu, J.; Allerson, C. R.; Berdeja, A.; Swayze, E. E. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 1122 – 1125.

[8] Seth, P. P.; Siwkowski, A.; Allerson, C. R.; Vasquez, G.; Lee, S.; Prakash, T. P.; Kinberger, G.; Migawa, M. T.; Gaus, H.; Bhat, B.; Swayze, E. E. *Nucleic Acids Symp. Ser.* **2008**, *52*, 553 – 554.

[9] Vasquez, G. Patent WO 2009/143369 A2, 2009.

[10] Swayze, E. E.; Seth, P. P. Patent WO 2007/090071 A2, 2007.

[11] The 12 steps required to synthesise the common intermediate **4** are only counted once.

[12] Defects in the modular sugar core undergo net amplification when converted into oligonucleotide chains. For example, four nucleosides (A, G, C, U) derived from a common sugar (S) building block containing x% of an anomalous sugar impurity (S*) will each contain x% of the related nucleoside impurity (A*, G*, C*, U*). When multiple (n) nucleosides are coupled into an oligonucleotide sequence, there will be a family of oligonucleotide impurities containing one (unspecified) anomalous nucleoside

(A-G-C…)* present at a level of nx%. The exact identity, as well as the toxicology and efficacy, of specific impurities in this family require careful consideration.

[13] For a survey of methods for glycosidation, see: Merino, P. *Chemical Synthesis of Nucleoside Analogues*; Wiley: 2013.

[14] The SELECT criteria, representing Safety, Environmental, Legal, Economy, Control and Throughput are frequently used as criteria for successful scale-up, see: Butters, M.; Catterick, D.; Craig, A.; Curzons, A.; Dale, D.; Gillmore, A.; Green, S. P.; Marziano, I.; Sherlock, J.-P.; White, W. *Chem. Rev.* **2006**, *106*, 3002 – 3027.

[15] Koshkin, A. A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J. *Tetrahedron* **1998**, *54*, 3607 – 3630.

[16] The 6'-epimer of epoxide **10**, generated in 13 steps from **2**, has been used previously for the synthesis of nucleosides but is stereochemically-unsuitable for application to cEt structures, see:Enderlin, G.; Nielsen, P. *J. Org. Chem.* **2008**, *73*, 6891 – 6894.

[17] Turks, M.; Rodins, V. Rolava, E.; Ostrovskis, P.; Belyakov, S. *Carbohydr. Res.* **2013**, *375*, 5 – 15.

[18] Fukuyama, K.; Ohrui, H.; Kuwahara, S. *Org. Lett.* **2015**, *17*, 828 – 831.

[19] Szeja, W. *Carbohydr. Res.* **1986**, *158*, 245 – 248.

[20] Vorbruggen, H.; Ruh-Pohlenz, C. *Org. React. (Hoboken, NJ, U. S.)* **2000**, 55.

[21] Cipolla, L.; Airoldi, C.; Bini, D.; Gregori, M.; Marcelo, F.; Jiménez-Barbero, J.; Nicotra, F. *Eur. J. Org. Chem.* **2011**, 128 – 136.

[22] In keeping with the step-counting approach used previously (**SCHEME 1** and ref [11], the 8 steps to common intermediate **11** are only counted once.